

CLAIMS:

1. In a computer system that implements a search engine to identify descriptions that match a set of key words, a method of enhancing user input to improve discovery, the method comprising the computer-implemented steps of:
 - a. Analyzing the terminology usage within the database to identify the core vocabulary;
 - b. assisting in the identification of the predominate semantic structure;
 - c. recording the conceptual assignment, supplementary terms and contextual significance of the core vocabulary;
 - d. receiving a search query from a user, the search query including at least one query term;
 - e. supplementing the search query with semantic data associated with the input query term(s);
 - f. identifying database descriptions that are conceptually similar to the input;
 - g. ranking identified description based on their similarity to the input;
 - h. Presenting the similar entries to the user for subjective selection.
2. The method of claim1, wherein step (c) comprises generating a data structure which links key terms to other terms related to them within the context of the database as well as their contextual significance, based on their predominate semantic usage, and step (e) comprises accessing the data structure to add the related terms and their contextual significance to the query criteria.

3. The method of claim 1, wherein step (a) comprises the sub-steps of:

- (a1) creating a frequency distribution analysis of the words used in the database descriptions; and
- (a2) rank ordering the words in descending order of usage; and
- (a3) identifying the word where the second derivative of individual usage with respect to the cumulative number of words analyzed reaches its first local minimum; and
- (a4) identifying the set of words, from most used to the word identified in (a3), which compose the core vocabulary of the database.

4. The method of claim 1, step (b) comprises the sub-steps of:

- (b1) presenting a statistically valid sample of descriptions that contain the word for manual review; and
- (b2) presenting a template of common conceptual groupings for manual review; and
- (b3) manually identifying and recording a list of the conceptual groupings that predominate the semantic structure of the database descriptions and assigning an importance level to each grouping.

5. The method of claim 1, step (c) comprises the sub-steps of:

- (c1) for each term in the core vocabulary, presenting a statistically valid sample of its citations for manual review; and
- (c2) presenting the list of conceptual groupings developed in step (a7) for manual review; and
- (c3) manually assigning the term to a conceptual grouping.

OCT-16-00 MON 07:14 AM BECKER CAPITAL

FAX NO. 3035272799

P. 15/26

6. The method of claim 1, step (c) further comprising the sub-steps of:
 - (c4) preparing a lexicon for the database composed of a record of each term in the core vocabulary, its conceptual grouping as well as its importance category; and
 - (c5) appending to each record all other terms in the lexicon that share the primary term's conceptual grouping; and
 - (c6) manually reviewing each entry and judgmentally adding appropriate synonyms, anonyms and common misnomers; and
 - (c7) manually assigning each term in each entry a relationship to the primary term; and
 - (c8) assigning a significance factor to each term of each entry based on a look-up table matrix of grouping importance and term relationship.
7. The method of claim 1, wherein step (f) comprises generating a similarity index for each database description based on the query term(s) and their associated semantic data.
8. The method of claim 7, wherein step (f) comprises the sub-steps of:
 - (f1) creating a similarity index which is initially set at zero; and
 - (f2) for each term in the expanded query criteria, comparing it to the words in the database description; and
 - (f3) in the event of a word match, indexing the entry's similarity factor by the term's significance factor.
9. The method of claim 1, wherein step (f) further comprises the sub-steps of:
 - (f4) identifying for output entries that have a positive similarity index.